

Parallel kinetic Monte Carlo simulation framework incorporating accurate models of adsorbate lateral interactions

Jens Nielsen, Mayeul d'Avezac, James Hetherington, and Michail Stamatakis

Citation: *The Journal of Chemical Physics* **139**, 224706 (2013); doi: 10.1063/1.4840395

View online: <http://dx.doi.org/10.1063/1.4840395>

View Table of Contents: <http://scitation.aip.org/content/aip/journal/jcp/139/22?ver=pdfcov>

Published by the [AIP Publishing](#)



Re-register for Table of Content Alerts

Create a profile.



Sign up today!



Parallel kinetic Monte Carlo simulation framework incorporating accurate models of adsorbate lateral interactions

Jens Nielsen,¹ Mayeul d'Avezac,¹ James Hetherington,¹ and Michail Stamatakis^{2,a)}

¹Research Software Development Team, Research IT Services, University College London, Torrington Place, London WC1E 6BT, United Kingdom

²Department of Chemical Engineering, University College London, Torrington Place, London WC1E 7JE, United Kingdom

(Received 30 September 2013; accepted 20 November 2013; published online 12 December 2013)

Ab initio kinetic Monte Carlo (KMC) simulations have been successfully applied for over two decades to elucidate the underlying physico-chemical phenomena on the surfaces of heterogeneous catalysts. These simulations necessitate detailed knowledge of the kinetics of elementary reactions constituting the reaction mechanism, and the energetics of the species participating in the chemistry. The information about the energetics is encoded in the formation energies of gas and surface-bound species, and the lateral interactions between adsorbates on the catalytic surface, which can be modeled at different levels of detail. The majority of previous works accounted for only pairwise-additive first nearest-neighbor interactions. More recently, cluster-expansion Hamiltonians incorporating long-range interactions and many-body terms have been used for detailed estimations of catalytic rate [C. Wu, D. J. Schmidt, C. Wolverton, and W. F. Schneider, *J. Catal.* **286**, 88 (2012)]. In view of the increasing interest in accurate predictions of catalytic performance, there is a need for general-purpose KMC approaches incorporating detailed cluster expansion models for the adlayer energetics. We have addressed this need by building on the previously introduced graph-theoretical KMC framework, and we have developed Zacros, a FORTRAN2003 KMC package for simulating catalytic chemistries. To tackle the high computational cost in the presence of long-range interactions we introduce parallelization with OpenMP. We further benchmark our framework by simulating a KMC analogue of the NO oxidation system established by Schneider and co-workers [*J. Catal.* **286**, 88 (2012)]. We show that taking into account only first nearest-neighbor interactions may lead to large errors in the prediction of the catalytic rate, whereas for accurate estimates thereof, one needs to include long-range terms in the cluster expansion. © 2013 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution 3.0 Unported License. [<http://dx.doi.org/10.1063/1.4840395>]

I. INTRODUCTION

The origins of temporal investigations of processes on lattices can be sought in the seminal work by Glauber, who analyzed a Markov process for spin flips in the Ising Model, back in 1963.¹ The subsequent development of kinetic Monte Carlo approaches^{2,3} enabled the study of complex time-dependent phenomena in the Ising model such as metastability and dynamic critical phenomena (see for instance Refs. 4–7). Application of kinetic Monte Carlo (KMC) to catalysis on surfaces was pioneered by Ziff, Gulari, and Barshad,⁸ who investigated kinetic phase transitions on a model for CO oxidation in 1986. Subsequent studies focused on a variety of temporal phenomena in catalytic systems, such as bistable responses, noise-induced transitions, as well as oscillatory behavior.^{9–15}

Since these early studies, significant developments have taken place. Over the past two decades KMC has evolved

to enable the coupling with *ab initio* calculations for a first-principles based simulation of chemistries on catalytic surfaces^{16–31} (for reviews, see Refs. 32–34). In these first-principles KMC frameworks the reaction energies and activation barriers are typically obtained from density functional theory (DFT) calculations, and the kinetic parameters are calculated by employing transition state theory (TST) approximations. Thus, within the level of accuracy of DFT and to the extent of validity of TST, the quality of the predictions obtained by KMC depends on “how well” the *ab initio* data have been incorporated into the simulation.

In particular, attractive or repulsive interactions between adsorbates on the catalytic surface have been shown to affect the rates of elementary reactions.^{16,35} Early first-principles KMC frameworks used DFT and bond-order conservation (BOC) methods to account for such effects.^{16,17} In some instances, such interactions were neglected altogether or modeled by pairwise additive nearest neighbor contributions, due to large computational expense needed for implementing more accurate models.³⁶ The most general such models consist of cluster-expansion (CE) Hamiltonians, which can accommodate any level of accuracy by taking into account

^{a)} Author to whom correspondence should be addressed. Electronic mail: m.stamatakis@ucl.ac.uk. Tel.: +44-203-108-1128.

long-range and many-body contributions to the total energy. This approach dates back to the 1980s,³⁷ and has been applied extensively to study the thermodynamics of adlayer structures.^{38–45} Moreover, CEs have been employed to investigate adsorption/desorption dynamics^{46–48} and the diffusion of adatoms on metal systems.^{49,50}

Only recently however, have cluster expansion Hamiltonians been applied for the estimation of catalytic rates.⁵¹ In an elegant study, Schneider and co-workers investigated the kinetics of NO oxidation on Pt(111) by means of equilibrium Monte Carlo (MC) calculations and appropriate averaging of the microscopic kinetic rates over the lattice. The MC calculations employed a cluster expansion for O on Pt(111)⁴⁵ and yielded equilibrium structures of the oxygen adlayer. The dissociative adsorption energy of O₂ was then calculated for each pair of empty sites (using the cluster expansion) and mapped to an activation energy for O₂ dissociation thereon via a Brønsted-Evans-Polanyi (BEP) relationship. Thus, a distribution of activation energies was obtained, from which the average rate of O₂ dissociation was evaluated. To make the connection with NO oxidation, it was assumed that the NO/NO₂ mixture was at equilibrium with adsorbed oxygen, thereby computing the chemical potential of O which is the input to the aforementioned MC calculations.

In view of the increasing interest in accurate predictions of catalytic performance, long-range lateral interactions and possibly many-body contributions are expected to form an integral part of kinetic models of catalysts. We have thus developed a general-purpose KMC framework that incorporates detailed cluster expansion models for the adlayer energetics, building on the previously introduced graph-theoretical KMC approach.⁵² Using this framework, one can define a cluster expansion model for the adlayer energetics, along with the forward activation energies of the elementary reactions for the chemistry under investigation. The reverse activation energies are computed through linear BEP relations thereby ensuring that the simulation exhibits microscopic reversibility. To tackle the high cost of computing the energetics in the presence of long-range interactions we introduce parallelization with OpenMP, and benchmark the performance of the framework by simulating a KMC analogue for the NO oxidation model established by Schneider and co-workers.⁵¹

The rest of the paper is organized as follows. In Sec. II we present a brief overview of the graph-theoretical KMC framework, and present our current development work incorporating cluster expansion Hamiltonians in the framework. Subsequently, we discuss the setting up of a benchmark model for NO oxidation on Pt(111) based on the cluster expansions developed by Schneider and co-workers.^{45,51} In Sec. III, we validate our computational framework and investigate the behavior of the model under different operating conditions. We further discuss the performance and scalability of the parallel code. In Sec. IV, we highlight our contributions and discuss the significance of this work.

II. METHODOLOGY

The components and procedures of the graph-theoretical KMC implementation are discussed in Ref. 52, whereas

Refs. 53–56 demonstrate the use of the method in gaining a fundamental understanding of surface processes in a variety of catalytic systems. For a general overview of KMC approaches and applications in heterogeneous catalysis the reader is referred to Ref. 55. Here, we will briefly review the framework and present in detail the calculation of the energetics which is the focus of the present work.

A. Overview of graph-theoretical KMC

The workload of the KMC algorithm consists of simulating a sequence of elementary events, such as adsorption, desorption, diffusion, reaction, which change the configuration of a lattice representing the catalytic surface. In the graph-theoretical KMC, the lattice, adsorbate configurations, elementary reactions, as well as energetic interaction contributions, are all represented by graphs. The input to a simulation consists of parameters controlling the behavior of the program (for instance, sampling times, stopping criteria, etc.), and the components specifying the physics of the system under investigation. For the latter, one has to provide the conditions (T, P, and gas phase composition), a lattice structure, an initial configuration of the lattice, an energetics model containing at minimum the energies of gas species and the binding configurations of adsorbates, and finally, a mechanistic model containing a list of possible elementary events. Note that in our previous work, no energetic model was specified. The code could only handle pairwise additive interactions and microscopic reversibility was manually implemented by correctly listing all interactions between neighboring species and the initial and final state species of each elementary event. In the current work, we introduce cluster-expansion based models for a general and thermodynamically consistent KMC simulation scheme.

As in previous KMC schemes, a simulation is initialized with the given lattice configuration (possibly an empty lattice), and a queue is generated containing all the possible lattice processes for the given configuration. For instance, if the lattice is initialized with one adsorbed molecule, the possible processes could be: adsorption on all empty sites, desorption of the adsorbate, or diffusional hops to each of the neighboring empty sites. The queue just mentioned also stores randomly generated times in which each process can occur. Under constant conditions, each of these random times follows an exponential distribution with a rate parameter equal to the kinetic constant of that process.⁵⁷

After initialization, the algorithm enters a loop at each step of which the most imminent process in the queue (the one with the smallest time of occurrence) is executed. Thus, the reactants are removed from the lattice along with the processes they participate in, and the products are added to the lattice. The latter operation also involves detecting all the possible processes in which the newly added adsorbates can participate and including them in the queue of lattice processes. This detection entails searching the neighborhood of each newly added adsorbate to identify patterns matching an elementary event, e.g., desorption or diffusion. Note that for each of these patterns identified, a rate constant has to be

calculated. Thus, one needs to know the activation energy, which in general depends on the local environment due to the lateral interactions. Moreover, the removal of adsorbed reactants from the lattice eliminates energetic interactions, and the addition of adsorbed reaction products introduces new energetic interactions which affect the activation energies and consequently the rates of existing processes. Computing the activation energies can be computationally intensive if the energetics' model contains long-range interactions. Thus, we implemented parallel computing in order to improve the efficiency of these calculations.

The procedures just described occur at every KMC step. By simulating a long sequence of such steps, a KMC trajectory (realization) is generated which can be subsequently post-processed to yield a variety of observables. For instance, counting the number of adsorbates on each site type yields surface coverages from which the most abundant surface intermediate can be found. Counting the number of reactant molecules consumed per site per time allows one to calculate the turnover frequency of the overall reaction. Moreover, the rates of each of the elementary reactions can be calculated by counting the occurrences thereof per unit time. Analysis of this data can reveal the most dominant pathway and sensitivity analysis can elucidate the processes limiting the overall reaction rate, namely the rate determining steps (RDS).^{58,59} KMC simulations thus provide a wealth of information that can be used to elucidate the underlying molecular phenomena giving rise to catalysis.

B. Implementation of detailed models of lateral interactions

Adsorbate lateral interactions have attracted the interest of experimental and theoretical studies for many years.⁶⁰ In catalysis, interactions between reactant species and spectator species in the neighborhood of a microscopic event can give rise to spatial heterogeneity in chemical reactivity.⁶¹ It is thus important to model such effects in order to enable the accurate prediction of catalytic performance.

To this end, it is necessary to formulate a model Hamiltonian, in order to capture the energetics of the adsorbate overlayer. For lattice systems, the most general such Hamiltonian is given by the so-called cluster expansion.³⁷ Mathematically, for a specified lattice and a given function that maps the occupancy of the lattice points to a real number, for instance, the energy for a given configuration of adsorbates, one can in principle construct a cluster expansion that represents this function exactly. Of course, for practical purposes one usually truncates the expansion, which introduces error. The basic idea behind this model is to decompose the energy of a lattice configuration into single-body, two-body, and many-body interaction terms represented by clusters (also known as figures). From a technical standpoint, the state of each site (which in our case represents which adsorbate is bound thereon) is given by a spin variable. One defines a set of basis functions for each site, consisting for instance of Chebyshev polynomials if an orthonormal basis is desired. Then, a set of cluster functions are defined as all the possible products of basis functions for each site in the cluster. To compute the en-

ergy of the lattice, one evaluates the linear combination of all cluster functions appropriately weighted with the contribution of each one of them to the overall energy of the system.³⁷ The method has recently been expanded to accommodate different site types.⁶²

To implement a general cluster expansion Hamiltonian in the graph theoretical KMC, we adopt a slightly different formulation which is based on representing each cluster as a graph pattern, and counting the occurrences of that pattern by solving subgraph isomorphism problems. This makes it possible to accommodate patterns in which multidentate species participate. In the graph-theoretical KMC, the state of a single site σ_i is given by a 3-element vector,⁵²

$$\sigma_i \in \{1, 2, \dots, S_L\} \times \{0, 1, \dots, N_S\} \times \{1, 2, \dots, \text{maxdent}\} \\ \forall i \in \{1, \dots, S_L\}, \quad (1)$$

where S_L denotes the number of sites on the lattice, N_S the number of species, maxdent is the maximum number of sites that any species can occupy, for instance if species A binds to one site (monodentate) and species B binds to two (bidentate), $\text{maxdent} = 2$. In vector σ_i , the first element gives the index of the entity bound to that site; this is done in order to be able to tell the sites occupied by multidentate species. If only monodentate species are present, the maximum number of bound entities is S_L ; yet, in cases where multidentate species are bound on the lattice, this first element may range up to values less than S_L . The second element gives the species type, and the third element gives the dentate with which a species is bound to site i . Thus, the state of the overall lattice is given by an $S_L \times 3$ array σ ,⁵²

$$\sigma = \{\{\sigma_{i,j}\}_{j=1}^3\}_{i=1}^{S_L}. \quad (2)$$

In this setting, a graph pattern representing a cluster in the expansion is defined as

$$C_k = (\Xi_k, \mathcal{E}_k), \quad (3)$$

where Ξ_k represents the set of sites in cluster k (vertexes of the graph pattern), and \mathcal{E}_k the neighboring relations between these sites (edges of the graph). Each of the sites just noted can be assigned one out of the possible S_T site types that may exist on the lattice. Thus, if cluster k involves $S_{C,k}$ sites, the type of site k will be

$$\xi_{k,i} \in \{0, 1, 2, \dots, S_T\}, \quad \forall i \in \Xi_k = \{1, 2, \dots, S_{C,k}\}, \quad (4)$$

where the 0 means that a site k may be of any permitted type. In addition, the coverage pattern of cluster k can be defined by the states of all sites in the graph pattern in line with Eqs. (1) and (2),

$$\sigma_{k,i} \in \{\{1, 2, \dots, S_{C,k}\} \times \{0, 1, \dots, N_S\} \\ \times \{1, 2, \dots, \text{maxdent}(\mathbf{d})\}, \&\} \quad \forall i \in \Xi_k. \quad (5)$$

The above definition allows for the state of site i to be left unspecified, in which case it is denoted by the ampersand symbol (&). This is useful when defining long range interaction patterns in which the intermediate sites can be vacant or occupied by any species. Finally, a set of geometric constraints can be defined by specifying the angles between certain edges

in the graph. Thus, for these constraints a set of three sites $\{s_1, s_2, s_3\}$ in the cluster is defined and the angle between the vectors (x_1-x_2, y_1-y_2) and (x_3-x_2, y_3-y_2) is specified to the desired value,

$$\text{Angle}(s_{i,1}, s_{i,2}, s_{i,3}) = \varphi_i \in \{0, 2\pi\} \quad \forall i \in \{1, \dots, N_k^{\text{AC}}\}, \quad (6)$$

where the function Angle calculates the angle between the aforementioned vectors, φ_i is the desired angle for geometric constraint i , and N_k^{AC} gives the number of such constraints defined for cluster k .

Similar to the detection of lattice processes in the graph-theoretical KMC,⁵² an energetic contribution is a mapping between the vertex sets of \mathcal{C}_k (the subgraph) and $\mathcal{L} = (\mathcal{S}, \mathcal{E})$ (the “large” graph representing the lattice with sites \mathcal{S} and edges \mathcal{E}),

$$\mathcal{M} : \mathcal{E}_k \rightarrow \mathcal{S}. \quad (7)$$

This mapping has to satisfy the following conditions:

1. \mathcal{M} is a subgraph isomorphism, namely, every pair p, q of neighboring sites in pattern \mathcal{C}_k is mapped to a pair $\mathcal{M}(p), \mathcal{M}(q)$ of neighboring sites in the lattice:

$$\forall \begin{matrix} \{p, q\} \in \mathcal{C}_k \\ 1 \leq p \leq S_{C,k} \\ 1 \leq q \leq S_{C,k} \end{matrix} \Rightarrow \{\mathcal{M}(p), \mathcal{M}(q)\} \in \mathcal{E}. \quad (8)$$

2. Wherever specified, the angles between the specified edges of pattern \mathcal{C}_k are the same as those on the lattice:

$$\begin{aligned} \text{Angle}(\mathcal{M}(s_{i,1}), \mathcal{M}(s_{i,2}), \mathcal{M}(s_{i,3})) &= \varphi_i \\ \forall i &\in \{1, \dots, N_k^{\text{AC}}\}. \end{aligned} \quad (9)$$

3. Wherever specified, the types of sites of the pattern are the same as those on the lattice:

$$s_{\mathcal{M}(p),3} = \xi_{k,p} \quad \forall 1 \leq p \leq S_{C,k}. \quad (10)$$

4. There is a mapping between the cluster entities and the lattice entities:

$$\mathcal{F} : \{1, 2, \dots, N_{\text{CE}}\} \mapsto \{1, 2, \dots, N_E\}, \quad (11)$$

such that

$$\begin{aligned} \sigma_{\mathcal{M}(p)} &= (\sigma_{k,p,1}, \sigma_{k,p,2}, \mathcal{F}(\sigma_{k,p,3})) \\ \forall 1 \leq p \leq S_{C,k}, \text{ if } \sigma_{\mathcal{M}(p)} &\neq \&. \end{aligned} \quad (12)$$

Thus, the coverage patterns of the elementary step and the lattice match for sites with specified state.

Note that there are as many such mappings \mathcal{M} for cluster k (Eq. (7)) as the number of occurrences of that cluster on the lattice. This allows for a direct enumeration of the possible energetic cluster contributions.

Let $N_{\text{CE}_k}(\sigma)$ denote the number of occurrences of cluster \mathcal{C}_k for the lattice configuration σ and ECI_k the effective cluster interaction parameter, which gives the contribution of this

cluster to the total energy. To avoid overestimating the contributions, we define the graph multiplicity of cluster k , GM_k , which can be thought of as a symmetry number for that elementary pattern and gives the number of times the exact same pattern will be counted. For instance, pairwise energetic interactions between adsorbates of the same type occupying the same site types will be counted twice; thus, $\text{GM} = 2$. Then, the total energy of the system will be given as

$$\mathcal{H}(\sigma) = \sum_{k=1}^{N_c} \frac{\text{ECI}_k}{\text{GM}_k} \cdot N_{\text{CE}_k}(\sigma). \quad (13)$$

Equation (13) is the general cluster expansion Hamiltonian giving the energy of the adsorbate overlayer for any configuration.

C. Computing reaction energies and activation barriers

In the graph-theoretical KMC, every elementary event (adsorption, desorption, diffusion, reaction) is represented as a graph pattern with specified initial and final coverages. Due to lateral interactions, the reaction energy depends not only on the coverages of the sites involved in that event, but also on the coverages of sites in the neighborhood thereof. The reaction energy is thus given as

$$\Delta E_{\text{rxn}}^k(\sigma) = \mathcal{H}(\sigma'(\sigma, k)) - \mathcal{H}(\sigma) + \Delta E_{\text{gas}}^k, \quad (14)$$

where σ and σ' refer to the initial and final coverages of the overall lattice, and ΔE_{gas} is the change in the energy of gas species; for instance, if adsorbed CO and O gave rise to $\text{CO}_2(\text{gas})$, then ΔE_{gas} would be equal to the energy of CO_2 , approximately -3.1 eV referenced to CO and $1/2\text{O}_2$ in the gas phase. In practice, one does not need to recalculate from scratch the energy of the overall lattice in the final state $\mathcal{H}(\sigma')$. Rather, after computing $\mathcal{H}(\sigma)$ the energetic contributions of the reactants are subtracted therefrom, the clusters in which the products can participate are detected and the corresponding contributions are added therein to obtain $\mathcal{H}(\sigma')$.

Given the reaction energy, the activation energies of the forward and reverse elementary steps must satisfy the following relation to ensure microscopic reversibility (see Figure 1):

$$\Delta E_{\text{rxn}}^k(\sigma) = E_{\text{fwd}}^{k,\ddagger}(\sigma) - E_{\text{rev}}^{k,\ddagger}(\sigma). \quad (15)$$

The forward activation barrier can be parameterized by means of a BEP relation and *ab initio* data.⁵¹ Thus, for a given configuration, the forward activation energy can be expressed by a linear relation that involves the activation energy at the zero coverage limit, and the difference between the reaction energies at the given configuration and the zero coverage limit:

$$\begin{aligned} E_{\text{fwd}}^{k,\ddagger}(\sigma) &= \max(0, \Delta E_{\text{rxn}}^k(\sigma), \\ &E_{\text{fwd},0}^{k,\ddagger} + \omega \cdot (\Delta E_{\text{rxn}}^k(\sigma) - \Delta E_{\text{rxn},0}^k)). \end{aligned} \quad (16)$$

In the equation above, $E_{\text{fwd},0}^{k,\ddagger}$ and $\Delta E_{\text{rxn},0}^k$ are the forward activation barrier and reaction energy at the zero coverage limit

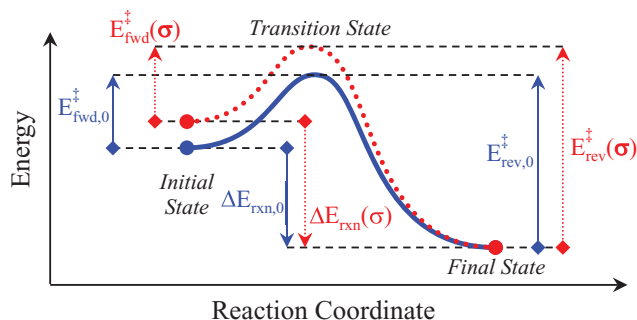


FIG. 1. Energy with respect to reaction coordinate showing the energetics- and kinetics-related quantities entering Eqs. (14)–(18). The blue solid path pertains to the reaction energy profile in the limit of zero coverage. For the red dotted path in this example we have assumed that repulsive interactions between the reactants and neighboring adsorbates raise the energy of the initial state.

(see Figure 1), and ω is a parameter referred to as the proximity factor.⁶³ The max operator ensures that the activation energy will be at least zero, or equal to the reaction energy if the latter is positive (otherwise the transition state energy would lie between the energies of reactants and products). In order to satisfy Eq. (15) the reverse activation energy has to be parameterized as

$$E_{rev}^{k,\ddagger}(\sigma) = \max(-\Delta E_{rxn}^k(\sigma), 0, E_{rev,0}^{k,\ddagger} - (1 - \omega) \cdot (\Delta E_{rxn}^k(\sigma) - \Delta E_{rxn,0}^k)), \quad (17)$$

where

$$E_{rev,0}^{k,\ddagger} = E_{fwd,0}^{k,\ddagger} - \Delta E_{rxn,0}^k. \quad (18)$$

Equations (16)–(18) allow us to compute the activation barriers in a thermodynamically consistent way, as prescribed by the cluster expansion Hamiltonian⁽¹³⁾ and the energetics of the gas species. The rate constant for reaction i can be then computed according to transition state theory as^{23,32,33,64,65}

$$k_{TST}^i = \frac{k_B \cdot T}{h} \cdot \frac{Q^\ddagger}{Q_R} \cdot \exp\left(-\frac{E^{i,\ddagger}(\sigma)}{k_B \cdot T}\right), \quad (19)$$

where Q^\ddagger and Q_R denote the quasi-partition functions of the transition state and reactants, respectively; h denotes Planck's constant, k_B Boltzmann's constant, and T the temperature. For details on the calculation of the quasi-partition functions and the rate constants for particular elementary events (for instance Eley-Rideal reactions) the reader can consult the supplementary material of Ref. 33.

Note that fitting the *ab initio*-computed activation energies to a BEP scaling relation introduces error. Our KMC approach however does not necessarily require the fitting of the activation energies to the BEP: if the activation energies for specific neighboring configurations are available, one can define several elementary reactions in which the arrangements of the spectator species (within the interaction range) and the corresponding activation energies are explicitly defined. However, this procedure is cumbersome from a practical standpoint, and thus, one resorts to an approximation, such as the one provided by the BEP relation. The latter, while not quan-

titative, is crucial for the qualitative understanding of the process of interest.

D. Algorithmic implementation

As implied by Eq. (13), computing the energy of the lattice necessitates detecting all energetic clusters and summing their contributions to the total energy. For bookkeeping purposes, the graph-theoretical KMC algorithm utilizes a data structure that stores the mapping information for each energetic cluster (namely, the lattice sites involved in the cluster), and also the inverse mapping that gives the indexes of the energetic clusters in which each adsorbed entity participates. This scheme is similar to that used for handling the elementary processes as described in Ref. 52. Thus, at every KMC step one knows the total lattice energy and all the energetic contributions it comprises. This information is used in calculating reaction energies and activation barriers as discussed in the following.

To calculate the activation energy for newly detected lattice processes (for instance an Eley-Rideal reaction step), the algorithm makes a temporary change in the state of the lattice, thereby executing that elementary step. In the final state, it detects the energetic clusters in which the products participate and computes the energy of the lattice from Eq. (13). The difference between final and initial state energies, calculated from Eq. (14), yields the reaction energy for the given configuration $\Delta E_{rxn}^k(\sigma)$. To find the reaction energy at the zero coverage limit, the algorithm performs a second computation, in which the clusters involving only the reactant entities are taken into account at the initial state, and similarly, the clusters involving only the product entities are considered in the final state, thereby neglecting interactions with neighboring adsorbates. The difference between the two energies thus computed gives $\Delta E_{rxn,0}^k$. Subsequently, the activation barriers can be computed via Eq. (16) or (17) depending whether the lattice process in discussion is a forward or a reverse step of an elementary event. Note that the user has to supply as input the forward activation energy at the zero coverage limit $E_{fwd,0}^{k,\ddagger}$, as well as the proximity factor ω . The reverse activation energy at zero coverage $E_{rev,0}^{k,\ddagger}$ is computed from Eq. (18).

The procedures just discussed take place whenever a new lattice process is detected in the course of a KMC simulation. The pseudocode for the whole algorithm has been given in our previous article; here we give a brief outline with specific focus on the calculation of energetics and kinetics.

0. Start

1. Define simulation lattice, conditions, participating species, elementary steps, energies of gas species, and cluster expansion Hamiltonian.
2. Initialize the lattice state and set the time clock to $t = 0$.
3. Detect and store all energetic cluster contributions in the energetics data-structure. Sum them up to obtain the lattice energy.
4. Detect all elementary events that can happen. For each such event, compute the activation energy and kinetic

constant, and generate a random time at which it will occur. Put these times in an event-queue.

5. While $t < t_{\text{final}}$
 - 5.a. Find the process μ that will occur next and update the time.
 - 5.b. Remove the reactants from the lattice. Also remove their energetic contributions from the energetics data-structure and the processes which they participate in from the event-queue.
 - 5.c. Add the products of process μ in the lattice. Detect their energetic cluster contributions and store them in the energetics data-structure. Update the lattice energy.
 - 5.d. For every product, detect the elementary events in which it can participate. For each of these events:
 - 5.d.i. Call subroutine *CalculateKineticConstant* to find the rate for the newly detected event.
 - 5.d.ii. Given the rate calculated from step 5.d.i, generate a random time for the occurrence of that event and store it in the event-queue.
 - 5.e. For each event within the neighborhood of the lattice process μ that just took place:
 - 5.e.i. Call subroutine *CalculateKineticConstant* to find the current rate in the presence of the new environment.
 - 5.e.ii. Given the rate calculated from step 5.e.i, generate a random time for the occurrence of that event and update the corresponding entry in the event-queue.
6. Repeat
7. **Terminate**

Subroutine *CalculateKineticConstant*

1. Calculate the initial state energy at the given configuration: from the energetics data-structure find the clusters in which reactants and neighboring adsorbates participate and sum up the contributions.
2. Calculate the initial state energy at the zero coverage limit: repeat the above calculation for clusters in which only reactants participate.
3. Make a temporary change in the state of the lattice to represent the final state.
4. Calculate the final state energy at the given configuration: detect the energetic clusters in which products and neighboring adsorbates participate and sum up the contributions.
5. Calculate the final state energy at the zero coverage limit: repeat the above calculation for clusters in which only products participate.
6. Revert the change in the state of the lattice done in step 3 of this subroutine.
7. Calculate the energy of reaction at the given configuration, $\Delta E_{\text{rxn}}^k(\sigma)$, and the zero coverage limit, $\Delta E_{\text{rxn},0}^k$, from Eq. (14).
8. Calculate the activation energy from Eq. (16) if this is a forward step of an elementary event, otherwise use Eqs. (17) and (18).
9. Calculate and return the kinetic constant from Eq. (19).

End Subroutine *CalculateKineticConstant*

E. Computational implementation and parallelization strategy

The above procedures were coded in our FORTRAN2003 software package, Zacros, which is currently available from the UCL online licensing portal, e-Lucid.⁶⁶ The package features an easy-to-learn keyword-based language for defining a simulation, and generates detailed output of the lattice state and energetics, reaction occurrence statistics, and numbers of molecules produced or consumed for each of the species participating in the chemistry. In addition, one can run simulations in debugging mode, during which the program provides an account of the changes in the key data structures during the simulation. This information can subsequently be used to efficiently troubleshoot a KMC simulation, if needed.

Using code profiling it was determined that the majority of the computational cost is due to the event time update operations in the loop of step 5.e. Thus, to improve performance, OpenMP parallelization was implemented over that loop. This scheme is applicable to shared memory architectures (for instance a single node in a computational cluster), and is based on the new reaction rate constants being calculated in parallel and stored into thread-private arrays. Subsequently, the random reaction occurrence times are generated serially and inserted into the event-queue, by looping over the threads and the processes assigned to every thread (double loop). This scheme produces simulation results which are identical to those generated by the serial version of the code.

F. Setting up a benchmark model

In a recent article, Schneider and co-workers, presented a set of cluster expansions for atomic oxygen adsorbed on Pt(111).⁴⁵ Careful error analysis of these expansions showed that for an accurate representation of the ground states of this system one needs to include up to 8th nearest neighbor pairwise additive interactions and at least two triplets (three-body contributions). Larger expansions were shown to be able to fit the DFT results quantitatively, but this essentially resulted in fitting the DFT error as well (overfitting).

In a subsequent study, Schneider and co-workers used the cluster expansion the strikes the best balance between accuracy and computational efficiency to study NO oxidation on Pt.⁵¹ Thus, they performed equilibrium MC simulations to calculate the distribution of activation energies and the overall rate of O₂ dissociation on the surface, given the chemical potential of O. The latter was calculated from the ratio of NO/NO₂ in the gas phase assuming that this mixture is in quasi-equilibrium with adsorbed O. Thus, the rate of NO oxidation was computed as one-half the rate of O₂ dissociative adsorption. In this work, associative desorption of O adatoms was neglected.

To benchmark our code we set up a KMC model for the NO oxidation using as input the energetic models (cluster expansions) and kinetic parameters by Schneider and co-workers.^{45,51,67}

TABLE I. Energetic contribution parameters (in meV) entering the lattice-gas Hamiltonian for O on Pt(111).

Figure	3-Figure CE	5-Figure CE	8-Figure CE	12-Figure CE
H_0/N_L	-27	45	9	12
Point	-1200	-1580	-1374	-1368
1NN	300	260	156	140
2NN		84	56	32
3NN		84	12	-16
4NN			28	28
5NN			28	32
6NN				12
7NN				8
8NN				12
1-1-3			64	56
1-2-3b				16

1. Energetics

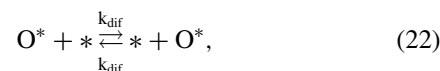
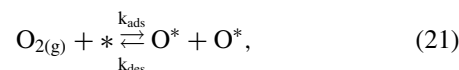
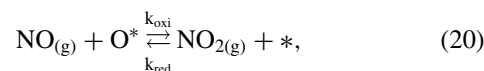
For developing the cluster expansion Hamiltonians capturing the adlayer energetics, Schmidt *et al.* used the Alloy Theoretic Automated Toolkit (ATAT)⁶⁸ coupled with the Vienna *Ab Initio* Simulation package (VASP).^{69–72} Plane-wave, DFT calculations were performed using the PW91 functional and the projector augmented wave (PAW) method.^{69,71–74} The slab models contained 4 layers of Pt (with only the bottom one fixed), a layer of oxygen adatoms and four vacuum layers, on $p(1 \times 1)$ up to $p(4 \times 4)$ supercells to investigate a variety of O coverages. This computational scheme was shown to incur a maximum error of 4 meV per adsorbed oxygen. The error due to uncertainty of the O_2 energy calculated from Generalized Gradient Approximation (GGA) was estimated to be 0.2 eV per adsorbed oxygen. For more information about the DFT methodology and its accuracy, the reader is referred to Ref. 45.

Schneider and co-workers^{45,51} thus developed four cluster expansion Hamiltonians of increasing accuracy by including 3, 5, 8, and 12 figures (clusters). The parameters reported in that work are for the Ising Hamiltonian formalism, which uses values of ± 1 for the state of a single site, denoting spin

up or down. For catalytic systems it seems more natural to adopt a lattice-gas Hamiltonian formulation in which the state of a single site takes values of 0 and 1, denoting a vacant or an occupied site accordingly. The mapping from the Ising parameters to the lattice-gas ones is discussed in detail in Sec. I of the supplementary material,⁷⁶ along with example calculations, and allows us to calculate the parameters of Table I from those of Table 2 in Ref. 45. These parameters correspond to the energetic contributions of the patterns shown in Figure 2. Note that the triplet 1-2-3b has a top site in the triangle sketched in Figure 2, and therefore, the only acceptable reflections and rotations are the ones drawn therein. Reflections in the horizontal axis would result in an hcp site inside the triangle and the figure would then be the 1-2-3a, which had not been included in this cluster expansion.^{45,51}

2. Kinetics

Oxygen adatoms are bound to fcc sites which is the only site type accounted for, in line with Refs. 45 and 51. In our model, the following reversible elementary events are explicitly considered:



where the last reaction denotes diffusion of O between neighboring sites. The graph patterns of these reactions as implemented in the graph-theoretical KMC are shown in Figure 3. In the following, we summarize the equations giving the rate constants of the above reactions. For a detailed discussion on how these equations were derived, the reader is referred to Sec. II of the supplementary material.⁷⁶

a. Oxidation of NO and reduction of NO₂. The 1st reaction (NO oxidation and reduction) was considered to be in quasi-equilibrium conditions by Wu *et al.*⁵¹ In the present work, the

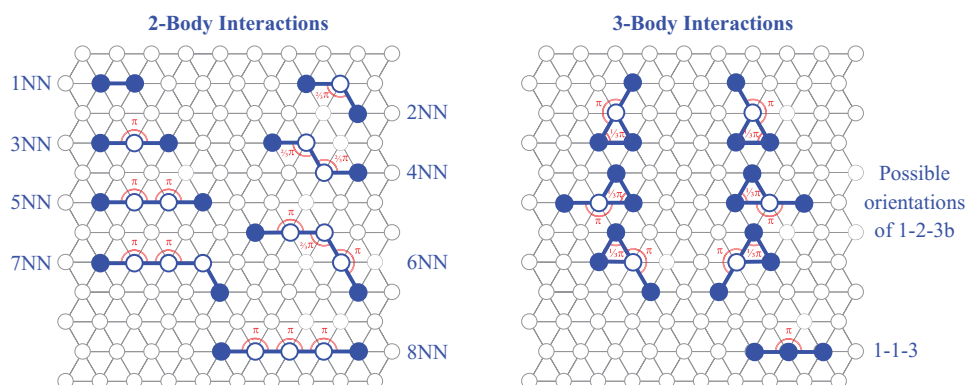


FIG. 2. Energetic interaction patterns taken into account in the cluster expansion Hamiltonians. Filled blue circles represent sites occupied by an oxygen adatom. Open blue circles represent sites that may or may not be occupied (their state is represented by the symbol $\&$ in Eq. (5)). The graph multiplicity for each of the pairwise interaction patterns is equal to 2. For the triplets, the graph multiplicity of the 1-1-3 pattern is also equal to 2, whereas the 1-2-3b patterns have multiplicities equal to 1.

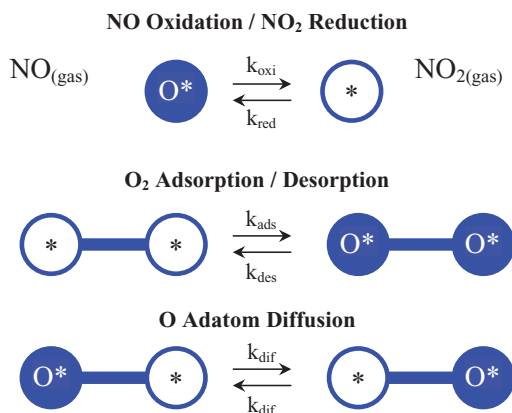


FIG. 3. Graph representations of the elementary events taken into account in the NO oxidation model.

quasi-equilibrium conditions are ensured by taking the forward and reverse steps of reaction (20) to be fast. Thus, we parameterize the kinetics constants thereof as follows:

$$k_{\text{red}} = A' \cdot P_{\text{NO}_2},$$

$$k_{\text{oxi}} = \left\{ \frac{A' \cdot P_{\text{NO}}}{Q_{\text{vib,O}}^{\text{surf}}} \cdot \exp\left(\frac{ZPE_{\text{O}}^{\text{surf}}}{k_B \cdot T}\right) \right\} \cdot \exp\left(-\frac{\Delta_f G_{\text{NO}_2}^0(T, 1\text{bar}) - FE_{\text{O}}^{\text{surf}} - \Delta_f G_{\text{NO}}^0(T, 1\text{bar})}{k_B \cdot T}\right), \quad (23)$$

where A' is a parameter that can be given a high value to ensure quasi-equilibration. The reduction and oxidation rates are proportional to the partial pressures of NO_2 and NO , P_{NO_2} , and P_{NO} , because of the participation of these gaseous species in the corresponding reactions. Moreover, $Q_{\text{vib,O}}^{\text{surf}}$ denotes the vibrational quasi-partition function of O adatoms (frequencies of 429, 380, 377 cm^{-1} , calculated from DFT and used in our calculations),⁴⁵ and $ZPE_{\text{O}}^{\text{surf}}$ the zero point energy thereof. Thus, the terms inside the curly brackets constitute the oxidation pre-exponential entering the KMC input. The activation energy at the zero coverage limit for this reaction is assumed to be negligible, and the reaction energy appearing in the exponent of the last term, is automatically handled by the KMC code. This reaction energy has three contributions, namely, the Gibbs free energies of formation of gas NO_2 , $\Delta_f G_{\text{NO}_2}^0(T, 1\text{bar})$, and NO , $\Delta_f G_{\text{NO}}^0(T, 1\text{bar})$, calculated from the formulas given in the NIST WebBook,⁷⁵ and the energy of formation of the O adlayer with respect to gas phase O_2 , $FE_{\text{O}}^{\text{surf}}$ calculated from the cluster expansions.

b. Adsorption of O₂. The rate of O_2 adsorption is given as

$$k_{\text{ads}} = \frac{\kappa_{\text{ads}}}{\sqrt{2 \cdot \pi \cdot m_{\text{O}_2} \cdot k_B \cdot T}} \cdot P_{\text{O}_2} \cdot \exp\left(-\frac{E_{\text{ads}}^{\ddagger}}{k_B \cdot T}\right)$$

$$= A_{\text{ads}} \cdot P_{\text{O}_2} \cdot \exp\left(-\frac{E_{\text{ads}}^{\ddagger}}{k_B \cdot T}\right). \quad (24)$$

In the above equation the orientational sticking factor κ_{ads} was arbitrarily taken to 1/6 as in Ref. 51. The activation energy for

the adsorption step follows the BEP relationship of Ref. 51:

$$E_{\text{ads}}^{\ddagger} = \max[0, FE_{2\text{O}}^{\text{surf}} + 2.12 \text{ eV}], \quad (25)$$

where $FE_{2\text{O}}^{\text{surf}}$ is the energy of formation of an O adatom pair on neighboring sites on the lattice. In order to implement this BEP relationship within the KMC framework (Eqs. (16) and (17)) we need to find the activation energy at the zero coverage limit and the proximity factor. The former is

$$E_{\text{ads,0}}^{\ddagger} = 2 \cdot ECI_{\text{O_Point}} + ECI_{\text{O_pair_1NN}} + 2.12 \text{ eV}, \quad (26)$$

where $ECI_{\text{O_Point}}$ is the 1-body term in the cluster expansion Hamiltonian, and $ECI_{\text{O_pair_1NN}}$ is the first nearest neighbor effective cluster interaction. Note that negative barriers are replaced by a zero value in Eq. (25). Finally, the proximity factor is equal to unity, which is expected as the transition state is a surface bound O_2 precursor.

c. Desorption of O₂. The rate constant of O_2 desorption is given as

$$k_{\text{des}} = \left\{ A_{\text{ads}} \cdot \frac{1}{(Q_{\text{vib,O}}^{\text{surf}})^2} \cdot \exp\left(\frac{2 \cdot ZPE_{\text{O}}^{\text{surf}}}{k_B \cdot T}\right) \right\} \cdot \exp\left(-\frac{-2 \cdot FE_{\text{O}}^{\text{surf}}}{k_B \cdot T}\right), \quad (27)$$

where the terms in curly brackets will be the pre-exponential entering the KMC input.

d. Diffusion of adsorbed O. An oxygen adatom can diffuse to a neighboring site with a rate that depends on the neighboring environment of the initial and the final site. Since the diffusion process is symmetric, we only consider the forward rate:

$$k_{\text{dif}} = \frac{k_B \cdot T}{h} \cdot \frac{Q^{\ddagger}}{Q_{\text{R}}} \cdot \exp\left(-\frac{E_{\text{dif}}^{\ddagger} - FE_{\text{O}}^{\text{surf}}}{k_B \cdot T}\right)$$

$$= A_{\text{dif}} \cdot \exp\left(-\frac{E_{\text{dif}}^{\ddagger} - FE_{\text{O}}^{\text{surf}}}{k_B \cdot T}\right). \quad (28)$$

To simulate the fast equilibration of the adsorbate overlayer, we adjusted the pre-exponentials for diffusion, A_{dif} , and NO oxidation/ NO_2 reduction, A' , such that: (i) the rates of these processes are at least 50 times larger than the O_2 dissociative adsorption (or O_2 associative desorption), and (ii) these processes are in partial equilibrium (forward divided by reverse rate is close to unity).

As an exemplar parameter set, Table II provides the pre-exponentials and activation energies (at the zero coverage limit) for the aforementioned elementary events at 480 K and $\ln(P_{\text{NO}_2}/P_{\text{NO}}) = -4.0$ for the 12-figure cluster expansion.

TABLE II. Exemplar kinetic parameters for the NO oxidation KMC model at 480 K and $\ln(P_{\text{NO}_2}/P_{\text{NO}}) = -4.0$ for the 12-figure CE.

Elementary event	Forward prefactor	Reverse prefactor	Proximity factor	Forward barrier (eV) at zero coverage
NO oxidation/NO ₂ reduction	$5.049 \times 10^{16} \text{ bar}^{-1} \text{ s}^{-1}$	$9.447 \times 10^{10} \text{ bar}^{-1} \text{ s}^{-1}$	0.00	0.000
O ₂ adsorption/desorption	$2.444 \times 10^6 \text{ bar}^{-1} \text{ s}^{-1}$	$6.980 \times 10^{17} \text{ s}^{-1}$	1.00	-0.476
O diffusion	$1.444 \times 10^4 \text{ s}^{-1}$	$1.444 \times 10^4 \text{ s}^{-1}$	0.50	0.100

III. RESULTS

A. Hull of ground states computed through simulated annealing

To validate the correct implementation of the cluster expansion Hamiltonian in the graph-theoretical KMC algorithm, we first computed the convex hull of the ground states for O on Pt(111). For each of these calculations, only diffusional hops were simulated at a given coverage. The lattice was initialized with a configuration of oxygen adatoms placed randomly on the lattice, such that the probability of occupancy of any site was equal to the coverage. Such configurations obviously have a high energetic cost, as they are far from the ordered adlayer structures observed at equilibrium. Subsequently, the system was allowed to relax while reducing the temperature linearly in time, thereby performing a simulated annealing calculation. After the system relaxed to a final state at a low temperature, we performed a second simulated annealing step, by restarting the system from that final state and the former initial temperature, and letting it relax again. This simulation setup ensures that the configuration minimizing the energy (ground state) will be found if the cooling rate is low enough.

The results of these calculations are shown in Figure 4. Panel (a) shows the temporal evolution of energy for a simulated annealing calculation on a 42×42 lattice (1764 sites) at a coverage equal to 0.35 ML. The energy is observed to drop quickly at initial times, as the adlayer relaxes. Moreover, as the temperature decreases linearly in time, the energy fluctuations are shown to drop, until no more events can happen. Figure 4(b) displays the convex hull of ground states, with some prominent structures noted thereon. These results are in agreement with the previously published ones by Schneider and co-workers,⁴⁵ thereby validating our methodology.

B. Dynamic simulations of NO oxidation and NO₂ reduction

To further test our methodology against the previously published results,⁵¹ we simulated the full dynamic model (reactions (20)–(22), Figure 3) containing, in addition to O diffusion, NO oxidation, NO₂ reduction, and O₂ dissociative adsorption. The reverse process of the latter was neglected for these comparisons, so that the latter are performed on an equal footing. We will come back to this point later.

The results of these simulations are summarized in Figure 5. Panels (a) and (b) show transients of the oxygen coverage and the number of NO₂ molecules produced or consumed per site. For this simulation, the lattice was initialized with a coverage of 0.25, which is lower than the stationary coverage of approximately 0.31. Thus, NO₂ reduction happens rapidly at initial times leaving oxygen adatoms on the surface. After reaction (20) has reached quasi-equilibrium, O₂ dissociation and slow NO oxidation are observed. A statistical analysis of the event occurrence is shown in the bar graph of Figure 5(c), which depicts the forward and reverse rates for each of the elementary steps taken into account in the model. It is evident that NO oxidation and NO₂ reduction are in partial equilibrium. O diffusion is observed to be in equilibrium as well. Since the diffusion step is symmetric with respect to the renumbering of the two sites, the observed partial equilibrium indicates that the energetic interactions of neighboring adsorbates have been correctly implemented in the forward and reverse diffusion rates. Finally, Figure 5(d) shows the turnover frequencies (TOFs) calculated by counting the NO₂ molecules produced per site per time after discarding the initial transient (see Figure 5(a)). Higher NO₂/NO ratios result in higher coverages of oxygen on the surface and thus TOFs. The latter are in excellent agreement with previously published results⁵¹ even though they have been computed with different methodologies.

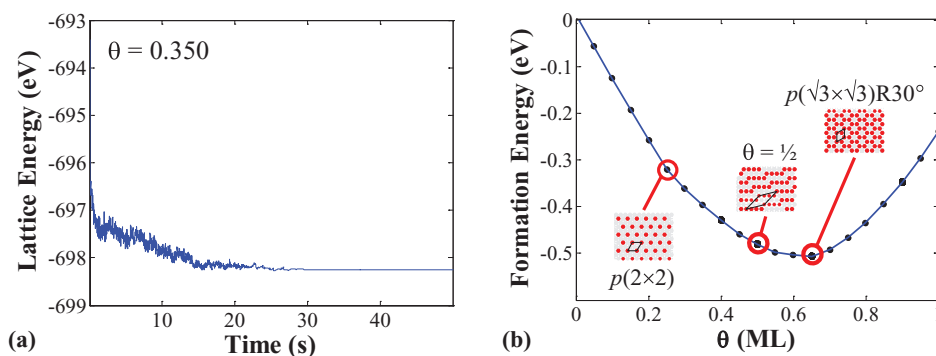


FIG. 4. Results of the simulated annealing calculations. (a) Energy versus time for $\theta = 0.35$, $T_{\text{ini}} = 100.5 \text{ K}$, $T_{\text{ramp}} = -2 \text{ K/s}$. The energy fluctuations drop over time and the ground state configuration is reached. (b) Convex hull of ground states. Prominent ground state structures are noted.

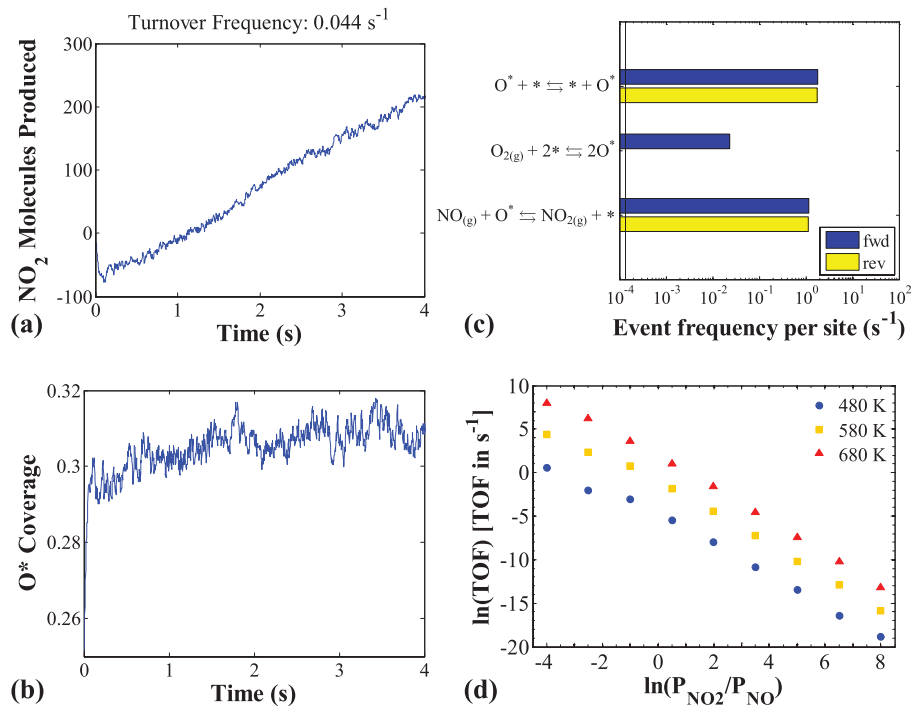


FIG. 5. Simulation results for the NO oxidation model with O_2 associative desorption neglected. (a) and (b) Number of NO_2 molecules produced or consumed per site and oxygen coverage, respectively. (c) Statistics of elementary event occurrence, showing the number of events per site per time for the forward and reverse steps. (d) Turnover frequencies (TOFs) for a range of ratios of NO_2/NO partial pressures. Each TOF was calculated from the slope of a plot similar to that of panel (a), after filtering out the initial transient. $P = 1$ atm and $y_{O_2} = 0.1$ for all simulations.

As previously mentioned, the simulations of Figure 5 neglect O_2 associative desorption for the comparison with published results to be valid. Moreover, from a physics standpoint, practical NO oxidation occurs at conditions at which O_2 desorption is not important. For high NO_2/NO ratios, however, it is expected that the rate of this step will be significant and dominate the mechanism, resulting in NO_2 reduction. One can actually compute the NO_2/NO ratio for which the overall chemistry will be at equilibrium given the molar fraction of O_2 :

$$\frac{(y_{NO_2})^2}{(y_{NO})^2 \cdot y_{O_2}} \cdot \frac{1 \text{ bar}}{P} = \exp\left(-\frac{\Delta G_{\text{overall}}^{\circ}(T, 1 \text{ bar})}{k_B \cdot T}\right) \quad (29)$$

and therefore,

$$\log\left(\frac{y_{NO_2}}{y_{NO}}\right) = \frac{1}{2} \cdot \log\left(y_{O_2} \cdot \frac{P}{1 \text{ bar}}\right) - \frac{\Delta G_{\text{overall}}^{\circ}(T, 1 \text{ bar})}{2 \cdot k_B \cdot T}, \quad (30)$$

where the change in the Gibbs free energy of the overall reaction $2NO + O_2 \leftrightarrow 2NO_2$ is

$$\begin{aligned} \Delta G_{\text{overall}}^{\circ}(T, 1 \text{ bar}) &= 2 \cdot G_{NO_2}^{\circ}(T, 1 \text{ bar}) - 2 \\ &\quad \cdot G_{NO}^{\circ}(T, 1 \text{ bar}) - G_{O_2}^{\circ}(T, 1 \text{ bar}). \end{aligned} \quad (31)$$

Simulations in which the O_2 associative desorption is taken into account are shown in Figure 6. Panel (a) shows the TOFs for a range of conditions, with the critical NO_2 to NO ratio computed from Eq. (30) noted in the plot. We observed that the KMC simulations correctly capture the transition from

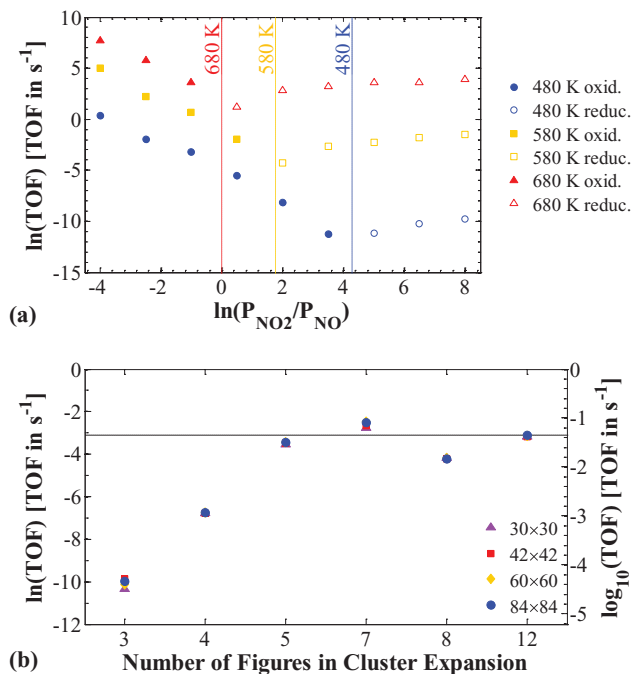


FIG. 6. Simulation results for the full NO oxidation model. (a) The vertical dashed lines correspond to the equilibrium NO_2/NO ratios for $y_{O_2} = 0.1$ and the three different temperatures (480 K, 580 K, and 680 K from right to left). On the left side of each line, NO oxidation is the net reaction, whereas on the right side, NO_2 reduction. (b) Turnover frequency predictions for six cluster expansions and progressively larger lattices. The conditions for these simulations were as follows: $T = 480$ K, $P = 1$ bar, $\log(y_{NO_2}/y_{NO}) = -1.0$. The lattice sizes used are noted in the legend, with the largest lattice (84×84) having 7056 sites.

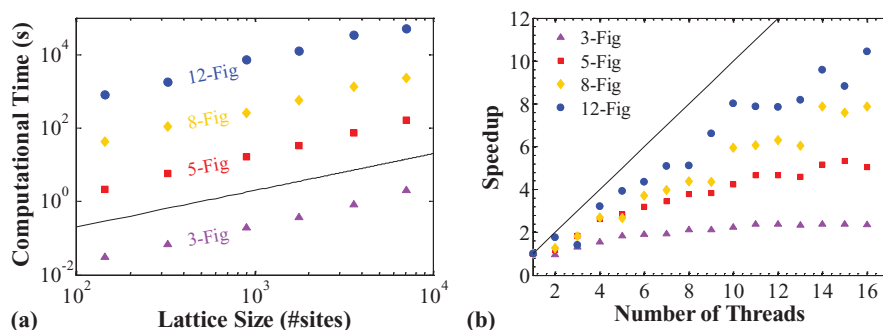


FIG. 7. Performance benchmarks: (a) Scaling of the computational time needed for propagating the system 1 simulated second, with respect to the lattice size. The black line is the graph of a linear equation: constant \times (number of sites). (b) Speedup with respect to number of threads. The speedup is defined as the ratio of the single-thread computational time with respect to the time of the multi-threaded simulation. The black line is the graph of a line with both slope and intercept equal to unity. Each point gives an average speedup computed from 10 runs.

NO oxidation on the left of the vertical lines, to NO₂ reduction on the right of the lines, where the number of gas NO₂ molecules decreases in time. This observation also verifies the correct implementation of the energetics model in the KMC simulation. It is interesting to observe that the net NO₂ reduction rate exhibits a negligible increase for higher NO₂/NO ratios. This can be attributed to the activation barrier for the associative desorption of oxygen being insensitive to the surface coverage thereof. Indeed, the proximity factor for the reversible O₂ adsorption (reaction (21)) is equal to 1 (see Table II), which means that the activation energy for oxygen desorption will be only weakly affected by lateral interactions. In the case of high enough coverages, for which O₂ dissociation is activated, the oxygen desorption barrier will be constant (see Eq. (17)). From a physical standpoint, this insensitivity means that the energetic interactions exerted by neighbors to the two O adatoms about to dissociate are similar to those exerted to the transition state (O₂ precursor). The slight increase in the NO₂ reduction rates for NO₂-richer mixtures is due to the higher numbers of neighboring O pairs on the surface, resulting from the higher oxygen coverages in this regime.

To evaluate the performance of the different CEs in predicting the TOFs we performed a set of calculations with progressively more accurate expansions. In addition to the four CEs already presented, we generated two more, capturing up to 2nd and 5th nearest-neighbor interactions (4- and 7-figure CEs, respectively). The latter CEs included at most 2-body terms and were generated by refitting configuration-energy data generated from the 12-figure CE. The configurations were randomly generated for a 10×10 lattice, and coverages uniformly distributed from 0 to 1. The parameters of these CEs are given in Sec. III of the supplementary material.⁷⁶

Figure 6(b) shows the results of simulations of the full NO oxidation NO₂ reduction system at $T = 480$ K and $\ln(P_{\text{NO}_2}/P_{\text{NO}}) = -1.0$. Simulations were performed for 4 different lattice sizes to exclude the possibility of error originating from small size effects. It is evident that the 3-figure CE, which includes only 1st nearest-neighbor pairwise additive interactions, errs by about 3 orders of magnitude. Inclu-

sion of 2nd nearest-neighbor interactions (4-figure CE), while providing some improvement, still underpredicts the TOF by 36-fold. The remaining cluster expansions (5-figure, 7-figure, 8-figure, and 12-figure CE) give similar results. These observations underscore the importance of including long-range interactions, for our purposes at least up to 3rd nearest-neighbor interactions. Taking into account multi-body terms, while certainly important for the correct prediction of ad-layer structures,^{43,45} does not seem to be critical for TOF estimations.

C. Performance benchmarks for parallel algorithm

The calculations involving long-range interactions are extremely computationally intensive: apart from the detection of energetic cluster contributions every time a new event is detected (steps 5.d of the pseudocode in Sec. II D), the algorithm also has to deal with the updates of existing lattice processes after the execution of an elementary event. To improve the performance of the algorithm, these updates were processed in parallel (see Sec. II E). To evaluate the performance of the algorithm we performed two benchmarks: (i) for a given number of threads we investigated how the computational time scales with respect to lattice size; (ii) for a given lattice size, we examined the speedup with respect to the number of threads. All benchmarks were performed on the computational cluster Legion@UCL on type U nodes with 16 cores. Furthermore, we allocated a full node independent of the thread count of the job to prevent any backfilling jobs from interfering with the benchmarking.

The results of the first benchmark are shown in Figure 7, which portrays the computational time required to propagate the system for 1 simulated second with respect to the number of sites in the lattice. We see that the computational time scales almost linearly with respect to the number of sites. This is intuitively expected as the number of lattice processes executed per unit time increases linearly with respect to the size of the system. Moreover, the computational expense of incorporating long range interactions into the KMC simulation becomes evident by the computational time for the 12-figure

CE being more than 4 orders of magnitude higher than that of the 3-figure CE. Improving the performance of the algorithm is thus of paramount importance.

Thus, Figure 7(b) shows the results of the second benchmark pertinent to the achieved speedup for the parallelized simulations. We observe that the improvement in the performance reaches a plateau as the number of threads increases. To understand why this is the case, we need to recall how our parallelization strategy works: the processes to be updated are first identified serially. Then, for each of these processes, the new activation energy and rate constant are computed; these calculations are the most intense and are performed in parallel. Subsequently, all new rate constants are collected from the different threads; this introduces a synchronization overhead. Finally, the new event occurrence times are computed serially and the queue of lattice processes is updated. If the number of threads is small with respect to the number of processes, the time required for the new rate constant calculations (in parallel) dominates, and the addition of more threads increases the speedup. On the contrary, if the number of threads is large, the synchronization when collecting the new rate constants is the “bottleneck” and the efficiency through parallelization is overshadowed by this synchronization overhead. Increasing the number of threads has a small effect on the speedup, which explains the plateaus observed.

This plateau behavior is reached “more quickly” (i.e., at lower number of threads) for the smaller CEs. This effect can be explained by the fact that the number of lattice processes in need of update after a KMC step increases for longer range interactions. In this case, the workload distribution via parallelization results in high efficiency, as opposed to when the number of updated processes is low, in which case the aforementioned communication overhead is significant.

IV. SUMMARY AND CONCLUSIONS

In view of the increasing interest in accurate predictions of reactions on catalytic surfaces, there is a need for general, flexible, and accurate kinetic modeling approaches. The graph-theoretical KMC method previously developed by Stamatakis and Vlachos⁵² allows for the direct incorporation of structurally complex elementary reactions computed by *ab initio* methods into KMC simulation. Thus, steric exclusions due to multidentate binding configurations, as well as intricate neighboring patterns of reactants and products were treated in a natural way within that formalism. The energetic interactions of the adlayer were constrained however to pairwise interactions and detailed balance was cumbersome to implement.

Cluster expansion Hamiltonians can be used to overcome both challenges with the construction of general models of adsorbate energetics, which may include many-body and long-range lateral interactions. Such Hamiltonians have been used extensively in the past to investigate the thermodynamics of adsorbate overlayers but there is only a limited number of studies focusing on the kinetics of reactions occurring in such environments. In this work, we incorporated general cluster expansion Hamiltonians in the graph-theoretical KMC framework, thereby enabling the detail treatment of adlayer ener-

getics in KMC simulation at any desired level of accuracy. Moreover, detailed balance is naturally incorporated into the framework, as the forward and reverse activation energies of a particular elementary step are contained to the thermodynamics of that step for the given neighboring environment. We have implemented this framework in Zacros, a computational KMC package written in FORTRAN2003, available through the UCL online licensing portal e-Lucid.⁶⁶ To improve the efficiency of the calculations for energetic models with long range interactions we explored parallelization with OpenMP.

To validate and benchmark our framework we set up a kinetic model for NO oxidation based on previous work by Schneider and co-workers.⁵¹ We performed simulated annealing calculations to verify that KMC correctly predicts the convex hull of ground states, and by means of dynamic simulations at a range of conditions we showed that our calculated TOFs agree with the ones found in the literature. Moreover, we demonstrated the effect of O₂ associative desorption, which was not accounted for in the original model. In this case, the NO₂/NO ratio in which the NO oxidation switches to NO₂ reduction was predicted by the KMC and was found to be in agreement with the ratio predicted purely by the thermodynamics of the gas phase reaction. We finally, showed that for accurate predictions of the TOF one needs to include long range interactions, while the 3-body terms are not critical.

It thus becomes evident that a KMC scheme incorporating detailed models for the adlayer energetics is a powerful tool for investigating surface kinetics. As massively parallel computer architectures become widely available, cluster expansion models with quantitative accuracy with respect to DFT calculations will be easily handled in KMC simulations. This will enable computational studies of the molecular-scale processes on catalytic surfaces at an unprecedented level of detail, unraveling new phenomena for chemistries of theoretical and practical importance.

ACKNOWLEDGMENTS

The research was partially supported by a new faculty start-up grant from the UCL Department of Chemical Engineering. The authors acknowledge the use of the UCL Research Software Development Service (RSD@UCL), as well as the UCL Legion High Performance Computing Facility (Legion@UCL) and associated support services in the completion of this work. M.S. is grateful to Professor William F. Schneider and Mr. Jason Bray for fruitful discussions.

¹R. J. Glauber, *J. Math. Phys.* **4**(2), 294 (1963).

²H. Müller-Krumbhaar and K. Binder, *J. Stat. Phys.* **8**(1), 1 (1973).

³A. B. Bortz, M. H. Kalos, and J. L. Lebowitz, *J. Comput. Phys.* **17** (1), 10 (1975).

⁴P. A. Flinn, *J. Stat. Phys.* **10**(1), 89 (1974).

⁵E. Stoll, K. Binder, and T. Schneider, *Phys. Rev. B* **8**(7), 3266 (1973).

⁶E. Stoll and T. Schneider, *Phys. Rev. A* **6**(1), 429 (1972).

⁷J. C. Angles d'Auriac, R. Maynard, and R. Rammal, *J. Stat. Phys.* **28**(2), 307 (1982).

⁸R. M. Ziff, E. Gulari, and Y. Barshad, *Phys. Rev. Lett.* **56**(24), 2553 (1986).

⁹M. Pineda, R. Imbühl, L. Schimansky-Geier, and Ch. Zülicke, *J. Chem. Phys.* **124**(4), 044701 (2006).

¹⁰M. Pineda, L. Schimansky-Geier, and R. Imbühl, *Phys. Rev. E* **75**(6), 061107 (2007).

- ¹¹V. P. Zhdanov and B. Kasemo, *Surf. Sci. Rep.* **20**(3), 113 (1994).
- ¹²V. P. Zhdanov and B. Kasemo, *Surf. Sci.* **405**(1), 27 (1998).
- ¹³V. P. Zhdanov and B. Kasemo, *Surf. Sci. Rep.* **39**(2–4), 25–104 (2000).
- ¹⁴D. J. Liu and J. W. Evans, *J. Chem. Phys.* **117**(15), 7319 (2002).
- ¹⁵D. G. Vlachos, L. D. Schmidt, and R. Aris, *Surf. Sci.* **249**(1–3), 248 (1991).
- ¹⁶M. Neurock and E. W. Hansen, *Comput. Chem. Eng.* **22**, S1045 (1998).
- ¹⁷E. W. Hansen and M. Neurock, *Chem. Eng. Sci.* **54**(15–16), 3411 (1999).
- ¹⁸E. W. Hansen and M. Neurock, *J. Catal.* **196**(2), 241 (2000).
- ¹⁹E. Hansen and M. Neurock, *J. Phys. Chem. B* **105**(38), 9218 (2001).
- ²⁰K. Reuter and M. Scheffler, *Phys. Rev. Lett.* **90**(4), 046103 (2003).
- ²¹K. Reuter, D. Frenkel, and M. Scheffler, *Phys. Rev. Lett.* **93**(11), 116105 (2004).
- ²²D. Mei, P. A. Sheth, M. Neurock, and C. M. Smith, *J. Catal.* **242**(1), 1 (2006).
- ²³K. Reuter and M. Scheffler, *Phys. Rev. B* **73** (4), 045433 (2006).
- ²⁴J. Rogal, K. Reuter, and M. Scheffler, *Phys. Rev. B* **77**(15), 155410 (2008).
- ²⁵S. Matera and K. Reuter, *Catal. Lett.* **133**(1–2), 156 (2009).
- ²⁶D. H. Mei, M. Neurock, and C. M. Smith, *J. Catal.* **268**(2), 181 (2009).
- ²⁷D. H. Mei, J. Du, and M. Neurock, *Ind. Eng. Chem. Res.* **49**(21), 10364 (2010).
- ²⁸H. A. Aleksandrov, L. V. Moskaleva, Z. J. Zhao, D. Basaran, Z. X. Chen, D. Mei, and N. Rösch, *J. Catal.* **285**(1), 187 (2012).
- ²⁹F. Hess, A. Farkas, A. P. Seitsonen, and H. Over, *J. Comput. Chem.* **33**(7), 757 (2012).
- ³⁰L. J. Xu, D. Mei, and G. Henkelman, *J. Chem. Phys.* **131**(24), 244520 (2009).
- ³¹P. Kratzer and M. Scheffler, *Phys. Rev. Lett.* **88**(3), 036102 (2002).
- ³²K. Reuter, “First-principles kinetic Monte Carlo simulations for heterogeneous catalysis: Concepts, status and frontiers,” in *Modeling Heterogeneous Catalytic Reactions: From the Molecular Process to the Technical System*, edited by O. Deutschmann (Wiley-VCH, Weinberg, 2010).
- ³³M. Stamatakis and D. G. Vlachos, *ACS Catal.* **2**(12), 2648 (2012).
- ³⁴M. K. Sabbe, M.-F. Reyniers, and K. Reuter, *Catal. Sci. Technol.* **2**(10), 2010 (2012).
- ³⁵M. Neurock, R. A. van Santen, W. Biemolt, and A. P. J. Jansen, *J. Am. Chem. Soc.* **116**(15), 6860 (1994).
- ³⁶K. Reuter, C. Stampfl, and M. Scheffler, “*Ab initio* thermodynamics and statistical mechanics of surface properties and functions,” in *Handbook of Materials Modeling. Part A: Methods* (Springer, Berlin, 2005), p. 149.
- ³⁷J. M. Sanchez, F. Ducastelle, and D. Gratias, *Physica A* **128**(1–2), 334 (1984).
- ³⁸R. Drautz, R. Singer, and M. Fähnle, *Phys. Rev. B* **67**(3), 035418 (2003).
- ³⁹H. R. Tang, A. Van der Ven, and B. L. Trout, *Phys. Rev. B* **70**(4), 045420 (2004).
- ⁴⁰H. R. Tang, A. Van der Ven, and B. L. Trout, *Mol. Phys.* **102**(3), 273 (2004).
- ⁴¹C. Stampfl, *Phase Transitions* **80**(4–5), 311 (2007).
- ⁴²C. Lazo and F. J. Keil, *Phys. Rev. B* **79**(24), 245418 (2009).
- ⁴³S. D. Miller and J. R. Kitchin, *Mol. Simul.* **35**(10–11), 920 (2009).
- ⁴⁴S. Piccinin and C. Stampfl, *Phys. Rev. B* **81**(15), 155427 (2010).
- ⁴⁵D. J. Schmidt, W. Chen, C. Wolverton, and W. F. Schneider, *J. Chem. Theory Comput.* **8** (1), 264 (2012).
- ⁴⁶J. S. McEwen and A. Eichler, *J. Chem. Phys.* **126**(9), 094701 (2007).
- ⁴⁷S. H. Payne, J. S. McEwen, H. J. Kreuzer, and D. Menzel, *Surf. Sci.* **594**(1–3), 240 (2005).
- ⁴⁸C. Stampfl, H. J. Kreuzer, S. H. Payne, H. Pfnür, and M. Scheffler, *Phys. Rev. Lett.* **83**(15), 2993 (1999).
- ⁴⁹T. Rehman, M. Jaipal, and A. Chatterjee, *J. Comput. Phys.* **243**, 244 (2013).
- ⁵⁰S. Verma, T. Rehman, and A. Chatterjee, *Surf. Sci.* **613**, 114 (2013).
- ⁵¹C. Wu, D. J. Schmidt, C. Wolverton, and W. F. Schneider, *J. Catal.* **286**, 88 (2012).
- ⁵²M. Stamatakis and D. G. Vlachos, *J. Chem. Phys.* **134**(21), 214115 (2011).
- ⁵³M. D. Marcinkowski, A. D. Jewell, M. Stamatakis, M. B. Boucher, E. A. Lewis, C. J. Murphy, G. Kyriakou, and E. C. H. Sykes, *Nature Mater.* **12**(6), 523 (2013).
- ⁵⁴M. Stamatakis, Y. Chen, and D. G. Vlachos, *J. Phys. Chem. C* **115**(50), 24750 (2011).
- ⁵⁵M. Stamatakis, M. Christiansen, D. G. Vlachos, and G. Mpourmpakis, *Nano Lett.* **12**(7), 3621 (2012).
- ⁵⁶W. Guo, M. Stamatakis, and D. G. Vlachos, *ACS Catal.* **3**(10), 2248 (2013).
- ⁵⁷K. A. Fichthorn and W. H. Weinberg, *J. Chem. Phys.* **95** (2), 1090 (1991).
- ⁵⁸C. T. Campbell, *J. Catal.* **204**(2), 520 (2001).
- ⁵⁹H. Meskine, S. Matera, M. Scheffler, K. Reuter, and H. Metiu, *Surf. Sci.* **603**(10–12), 1724 (2009).
- ⁶⁰T. Matsushima, *Surf. Sci.* **558**(1–3), 1 (2004).
- ⁶¹M. Saliccioli, M. Stamatakis, S. Caratzoulas, and D. G. Vlachos, *Chem. Eng. Sci.* **66**(19), 4319 (2011).
- ⁶²D. Lerch, O. Wieckhorst, L. Hammer, K. Heinz, and S. Müller, *Phys. Rev. B* **78**(12), 121405(R) (2008).
- ⁶³L. C. Grabow, A. A. Gokhale, S. T. Evans, J. A. Dumesic, and M. Mavrikakis, *J. Phys. Chem. C* **112**(12), 4608 (2008).
- ⁶⁴P. Hänggi, P. Talkner, and M. Borkovec, *Rev. Mod. Phys.* **62**, 251 (1990).
- ⁶⁵P. Pechukas, “Statistical approximations in collision theory,” in *Dynamics of Molecular Collisions. Part B*, edited by W. H. Miller (Plenum Press, New York, London, 1976).
- ⁶⁶See http://www.e-lucid.com/i/software/material_modelling/Zacros.html for Zacros: Advanced Lattice-KMC simulation Made Easy.
- ⁶⁷J. M. Bray and W. F. Schneider, “First-principles thermodynamic models in heterogeneous catalysis,” in *Computational Catalysis*, edited by A. Asthagiri and M. J. Janik (Royal Society of Chemistry, 2013).
- ⁶⁸A. van de Walle, M. Asta, and G. Ceder, *CALPHAD: Comput. Coupling Phase Diagrams Thermochem.* **26**(4), 539 (2002).
- ⁶⁹G. Kresse and J. Furthmüller, *Comp. Mater. Sci.* **6**(1), 15 (1996).
- ⁷⁰G. Kresse and J. Furthmüller, *Phys. Rev. B* **54**(16), 11169 (1996).
- ⁷¹G. Kresse and D. Joubert, *Phys. Rev. B* **59**(3), 1758 (1999).
- ⁷²G. Kresse, M. Marsman, and J. Furthmüller, Vienna *ab initio* simulation package (VASP): The guide, see <http://cms.mpi.univie.ac.at/vasp/vasp/vasp.html>.
- ⁷³J. P. Perdew and Y. Wang, *Phys. Rev. B* **45**(23), 13244 (1992).
- ⁷⁴P. E. Blöchl, *Phys. Rev. B* **50**(24), 17953 (1994).
- ⁷⁵NIST, *National Institute of Standards and Technology—Chemistry Web-Book*, see <http://webbook.nist.gov/>.
- ⁷⁶See supplementary material at <http://dx.doi.org/10.1063/1.4840395> for (I) Converting a single-species Ising Hamiltonian to a lattice-gas Hamiltonian, (II) Derivation of kinetic rate constants for NO oxidation model, and (III) Lattice-gas parameters for all cluster expansions.